Analysis of Cereal data

Carl James Schwarz

October 23, 2019

```
> # Links for Sweave
>
> # Information about Sweave is at: https://rpubs.com/YaRrr/SweaveIntro
> # Chunk options       https://yihui.name/knitr/options/
>
> #Here we open the libraries needed and source any other code as needed
>
> library(car)
> library(emmeans)
> library(ggplot2)
> library(gridExtra)
> library(knitr)
> library(Hmisc)
> library(emmeans)
> library(pander)
> library(plyr)
> library(readxl)
> library(reshape2)
> source('http://www.stat.sfu.ca/~cschwarz/Stat-650/Notes/MyPrograms/schwarz.functions.r')
> # Get the raw data and do editing. This won't appear in the report, but gives you
> # an audit trail of what you did to the raw data before creating the rest of the
> # template
>
> cereal <- read.csv('cereal.csv',
+                 header=TRUE, as.is=TRUE, strip.white=TRUE)
> # Define new variables and factors (for categorical variables). CHeck the structure of the
> cereal$shelfF <- factor(cereal$shelf)
```

1 Introduction

What is the relationship between the calories in a serving of breakfast cereal
and the placement of the cereal on the shelf in the supermarket?

2　Material and Methods

A sample of 77 cereals were sampled from a local grocery store and the nutritional information (e.g. number of grams of fat, protein, carbohydrates, etc.) and the number of calories per serving was extracted. The display shelf on which the cereal was stored was also recorded.

A single-factor Completely Randomized Design (CRD) analysis of variance was used to compare the mean number of calories of cereals from each display shelf. This was followed by a Tukey multiple comparison procedure to investigate among which shelves the mean may differ.

All computations were performed using R version 3.6.1 (2019-07-05).

3　Results

The data was screened for outliers and no unusual points were located.

Table 1 summarizes the calories per serving by shelf number. Figure 1 shows a graphical display of the calories per serving.

Figure 1: Calories by shelf height

There was no evidence (p=0.54) of a difference in the mean calories per serving among the shelf height. But there was evidence of a difference in the mean grams of sugar per serving among the shelf height (p=0). A bar graph (Figure 2) of the mean grams per serving is shown below. There was no evidence that the mean amount of sugar per serving varied between the *Low* and *High* shelves, but there was evidence that the mean of both shelves differed from that on the *Middle* shelf.

Figure 2: Comparing the mean sugar content among shelves

Table 1: Summary statistics by shelf location

Shelf	n	Mean calories	Min calories	Max calories	SD calories
Low	20	100.5	80	120	11.5
Middle	21	107.6	80	130	12.2
High	36	106.1	40	160	29.0

4   Summary

We found no evidence that the mean number of calories varied among shelve heights, but there was evidence the the mean amount of sugars did.